

RECEIVED

MAY 15 2002

TECH CENTER 1600/2900

ENTERED

1642

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/202,047A

DATE: 05/07/2002

TIME: 15:57:05

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw

```

3 <110> APPLICANT: ITOH, Kyogo
4     SHICHIJO, Shigeki
5     IMAI, Yasuhisa
7 <120> TITLE OF INVENTION: TUMOR ANTIGEN PROTEINS, GENES THEREFOR, AND TUMOR
8     ANTIGEN PEPTIDES
10 <130> FILE REFERENCE: 0020-4491P
12 <140> CURRENT APPLICATION NUMBER: 09/202,047A
13 <141> CURRENT FILING DATE: 1998-12-07
15 <160> NUMBER OF SEQ ID NOS: 2
17 <170> SOFTWARE: PatentIn Ver. 2.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 2527
21 <212> TYPE: DNA
22 <213> ORGANISM: Homo sapiens
24 <220> FEATURE:
25 <221> NAME/KEY: misc_feature
26 <222> LOCATION: (1)..(2527)
27 <223> OTHER INFORMATION: Strandedness: Double-stranded
29 <220> FEATURE:
30 <221> NAME/KEY: 5'UTR
31 <222> LOCATION: (1)..(38)
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (39)..(2438)
37 <220> FEATURE:
38 <221> NAME/KEY: 3'UTR
39 <222> LOCATION: (2439)..(2506)
41 <220> FEATURE:
42 <221> NAME/KEY: polyA_site
43 <222> LOCATION: (2507)..(2527)
45 <400> SEQUENCE: 1
46 ggttcggcgg cagccgggct cggagtggac gtgccact atg ggg tcg tcc aag aag 56
47                                     Met Gly Ser Ser Lys Lys
48                                     1           5
50 cat cgc gga gag aag gag gcg gcc ggg acg acg gcg gcg gcc ggc acc 104
51 His Arg Gly Glu Lys Glu Ala Ala Gly Thr Thr Ala Ala Gly Thr
52      10           15           20
54 ggg ggt gcc acc gag cag ccg ccg cgg cac cgg gaa cac aaa aaa cac 152
55 Gly Gly Ala Thr Glu Gln Pro Pro Arg His Arg Glu His Lys Lys His
56      25           30           35
58 aag cac cgg agt ggc ggc agt ggc ggt agc ggt ggc gaa cga cgg aag 200
59 Lys His Arg Ser Gly Gly Ser Gly Gly Ser Gly Gly Glu Arg Arg Lys
60      40           45           50

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/202,047A

DATE: 05/07/2002

TIME: 15:55:05

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw

62 cgg agc cgg gaa cgt ggg ggc gag cgc ggg agc ggg cgg cgc ggg gcc 248  
 63 Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly Ser Gly Arg Arg Gly Ala  
 64 55 60 65 70  
 66 gaa gct gag gcc cgg agc agc acg cac ggg cgg gag cgc agc cag gca 296  
 67 Glu Ala Glu Ala Arg Ser Ser Thr His Gly Arg Glu Arg Ser Gln Ala  
 68 75 80 85  
 70 gag ccc tcc gag cgg cgc gtg aag cgg gag aag cgc gat gac ggc tac 344  
 71 Glu Pro Ser Glu Arg Arg Val Lys Arg Glu Lys Arg Asp Asp Gly Tyr  
 72 90 95 100  
 74 gag gcc gct gcc agc tcc aaa act agc tca ggc gat gcc tcc tca ctc 392  
 75 Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser Gly Asp Ala Ser Ser Leu  
 76 105 110 115  
 78 agc atc gag gag act aac aaa ctc cgg gca aag ttg ggg ctg aaa ccc 440  
 79 Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala Lys Leu Gly Leu Lys Pro  
 80 120 125 130  
 82 ttg gag gtt aat gcc atc aag aag gag gcg ggc acc aag gag gag ccc 488  
 83 Leu Glu Val Asn Ala Ile Lys Lys Glu Ala Gly Thr Lys Glu Glu Pro  
 84 135 140 145 150  
 86 gtg aca gct gat gtc atc aac cct atg gcc ttg cga cag cga gag gag 536  
 87 Val Thr Ala Asp Val Ile Asn Pro Met Ala Leu Arg Gln Arg Glu Glu  
 88 155 160 165  
 90 ctg cgg gag aag ctg gcg gct gcc aag gag aag cgc ctg ctg aac caa 584  
 91 Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu Lys Arg Leu Leu Asn Gln  
 92 170 175 180  
 94 aag ctg ggg aag ata aag acc cta gga gag gat gac ccc tgg ctg gac 632  
 95 Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu Asp Asp Pro Trp Leu Asp  
 96 185 190 195  
 98 gac act gca gcc tgg atc gag agg agc cgg cag ctg cag aag gag aag 680  
 99 Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg Gln Leu Gln Lys Glu Lys  
 100 200 205 210  
 102 gac ctg gca gag aag agg gcc aag tta ctg gag gag atg gac caa gag 728  
 103 Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu Glu Glu Met Asp Gln Glu  
 104 215 220 225 230  
 106 ttt ggt gtc agc act ctg gtg gag gag gag ttc ggg cag agg cgg cag 776  
 107 Phe Gly Val Ser Thr Leu Val Glu Glu Glu Phe Gly Gln Arg Arg Gln  
 108 235 240 245  
 110 gac ctg tac agt gcc cgg gac ctg cag ggc ctc acc gtg gag cat gcc 824  
 111 Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly Leu Thr Val Glu His Ala  
 112 250 255 260  
 114 att gat tcc ttc cga gaa ggg gag aca atg att ctt acc ctc aag gac 872  
 115 Ile Asp Ser Phe Arg Glu Gly Glu Thr Met Ile Leu Thr Leu Lys Asp  
 116 265 270 275  
 118 aaa ggc gtg ctg cag gag gag gag gac gtg ctg gtg aac gtg aac ctg 920  
 119 Lys Gly Val Leu Gln Glu Glu Glu Asp Val Leu Val Asn Val Asn Leu  
 120 280 285 290  
 122 gtg gat aag gag cgg gca gag aaa aat gtg gag ctg cgg aag aag aag 968  
 123 Val Asp Lys Glu Arg Ala Glu Lys Asn Val Glu Leu Arg Lys Lys Lys  
 124 295 300 305 310  
 126 cct gac tac ctg ccc tat gcc gag gac gag agc gtg gac gac ctg gcg 1016

RECEIVED  
 JUN 03 2002  
 TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/202,047A

TIME: 15:57:05

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw

127	Pro	Asp	Tyr	Leu	Pro	Tyr	Ala	Glu	Asp	Glu	Ser	Val	Asp	Asp	Leu	Ala	
128					315					320					325		
130	cag	caa	aaa	cct	cgc	tct	atc	ctg	tcc	aag	tat	gac	gaa	gag	ctt	gaa	1064
131	Gln	Gln	Lys	Pro	Arg	Ser	Ile	Leu	Ser	Lys	Tyr	Asp	Glu	Glu	Leu	Glu	
132				330					335						340		
134	ggg	gag	cgg	cca	cat	tcc	ttc	cgc	ttg	gag	cag	ggc	ggc	acg	gct	gat	1112
135	Gly	Glu	Arg	Pro	His	Ser	Phe	Arg	Leu	Glu	Gln	Gly	Gly	Thr	Ala	Asp	
136			345					350					355				
138	ggc	ctg	cgg	gag	cgg	gag	ctg	gag	gag	atc	cgg	gcc	aag	ctg	cgg	ctg	1160
139	Gly	Leu	Arg	Glu	Arg	Glu	Leu	Glu	Glu	Ile	Arg	Ala	Lys	Leu	Arg	Leu	
140		360					365					370					
142	cag	gct	cag	tcc	ctg	agc	aca	gtg	ggg	ccc	cgg	ctg	gcc	tcc	gaa	tac	1208
143	Gln	Ala	Gln	Ser	Leu	Ser	Thr	Val	Gly	Pro	Arg	Leu	Ala	Ser	Glu	Tyr	
144	375					380				385					390		
146	ctc	acg	cct	gag	gag	atg	gtg	acc	ttt	aaa	aag	acc	aag	cgg	agg	gtg	1256
147	Leu	Thr	Pro	Glu	Glu	Met	Val	Thr	Phe	Lys	Lys	Thr	Lys	Arg	Arg	Val	
148				395					400					405			
150	aag	aaa	atc	cgc	aag	aag	gag	aag	gag	gta	gta	gtg	cgg	gca	gat	gac	1304
151	Lys	Lys	Ile	Arg	Lys	Lys	Glu	Lys	Glu	Val	Val	Val	Arg	Ala	Asp	Asp	
152				410					415					420			
154	ttg	ctg	cct	ctc	ggg	gac	cag	act	cag	gat	ggg	gac	ttt	ggt	tcc	aga	1352
155	Leu	Leu	Pro	Leu	Gly	Asp	Gln	Thr	Gln	Asp	Gly	Asp	Phe	Gly	Ser	Arg	
156			425					430				435					
158	ctg	cgg	gga	cgg	ggt	cgc	cgc	cga	gtg	tcc	gaa	gtg	gag	gag	gag	aag	1400
159	Leu	Arg	Gly	Arg	Gly	Arg	Arg	Val	Ser	Glu	Val	Glu	Glu	Glu	Glu	Lys	
160		440					445				450						
162	gag	cct	gtg	cct	cag	ccc	ctg	ccg	tcg	gac	gac	acc	cga	gtg	gag	aac	1448
163	Glu	Pro	Val	Pro	Gln	Pro	Leu	Pro	Ser	Asp	Asp	Thr	Arg	Val	Glu	Asn	
164	455				460					465					470		
166	atg	gac	atc	agt	gat	gag	gag	gaa	ggt	gga	gct	cca	ccg	ccg	ggg	tcc	1496
167	Met	Asp	Ile	Ser	Asp	Glu	Glu	Glu	Gly	Gly	Ala	Pro	Pro	Pro	Gly	Ser	
168				475					480					485			
170	ccg	cag	gtg	ctg	gag	gag	gac	gag	gcg	gag	ctg	gag	ctg	cag	aag	cag	1544
171	Pro	Gln	Val	Leu	Glu	Glu	Asp	Glu	Ala	Glu	Leu	Glu	Leu	Gln	Lys	Gln	
172			490					495					500				
174	ctg	gag	aag	gga	cgc	cgg	ctg	cga	cag	tta	cag	cag	cta	cag	cag	ctg	1592
175	Leu	Glu	Lys	Gly	Arg	Arg	Leu	Arg	Gln	Leu	Gln	Gln	Leu	Gln	Gln	Leu	
176		505					510				515						
178	cga	gac	agt	ggc	gag	aag	gtg	gtg	gag	att	gtg	aag	aag	ctg	gag	tct	1640
179	Arg	Asp	Ser	Gly	Glu	Lys	Val	Val	Glu	Ile	Val	Lys	Lys	Leu	Glu	Ser	
180		520					525				530						
182	cgc	cag	cgg	ggc	tgg	gag	gag	gat	gag	gat	ccc	gag	cgg	aag	ggg	gcc	1688
183	Arg	Gln	Arg	Gly	Trp	Glu	Glu	Asp	Glu	Asp	Pro	Glu	Arg	Lys	Gly	Ala	
184	535				540				545					550			
186	atc	gtg	ttc	aac	gcc	acg	tcc	gag	ttc	tgc	cgc	acc	ttg	ggg	gag	atc	1736
187	Ile	Val	Phe	Asn	Ala	Thr	Ser	Glu	Phe	Cys	Arg	Thr	Leu	Gly	Glu	Ile	
188			555					560				565					
190	ccc	acc	tac	ggg	ctg	gct	ggc	aat	cgc	gag	gag	cag	gag	gag	ctc	atg	1784
191	Pro	Thr	Tyr	Gly	Leu	Ala	Gly	Asn	Arg	Glu	Glu	Gln	Glu	Glu	Leu	Met	

## RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/202,047A

TIME: 15:57:05

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw

```

192          570          575          580
194 gac ttt gaa cgg gat gag gag cgc tca gcc aac ggt ggc tcc gaa tct 1832
195 Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala Asn Gly Gly Ser Glu Ser
196          585          590          595
198 gac ggg gag gag aac atc ggc tgg agc acg gtg aac ctg gac gag gag 1880
199 Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr Val Asn Leu Asp Glu Glu
200          600          605          610
202 aag cag cag cag gat ttc tct gct tcc tcc acc acc atc ctg gac gag 1928
203 Lys Gln Gln Gln Asp Phe Ser Ala Ser Ser Thr Thr Ile Leu Asp Glu
204          615          620          625          630
206 gaa ccg atc gtg aat agg ggg ctg gca gct gcc ctg ctc ctg tgt cag 1976
207 Glu Pro Ile Val Asn Arg Gly Leu Ala Ala Ala Leu Leu Leu Cys Gln
208          635          640          645
210 aac aaa ggg ctg ctg gag acc aca gtg cag aag gtg gcc cgg gtg aag 2024
211 Asn Lys Gly Leu Leu Glu Thr Thr Val Gln Lys Val Ala Arg Val Lys
212          650          655          660
214 gcc ccc aac aag tcg ctg ccc tca gcc gtg tac tgc atc gag gat aag 2072
215 Ala Pro Asn Lys Ser Leu Pro Ser Ala Val Tyr Cys Ile Glu Asp Lys
216          665          670          675
218 atg gcc atc gat gac aag tac agc cgg agg gag gaa tac cga ggc ttc 2120
219 Met Ala Ile Asp Asp Lys Tyr Ser Arg Arg Glu Glu Tyr Arg Gly Phe
220          680          685          690
222 aca cag gac ttc aag gag aag gac ggc tac aaa ccc gac gtt aag atc 2168
223 Thr Gln Asp Phe Lys Glu Lys Asp Gly Tyr Lys Pro Asp Val Lys Ile
224          695          700          705          710
226 gaa tac gtg gat gag acg ggc cgg aaa ctc aca ccc aag gag gct ttc 2216
227 Glu Tyr Val Asp Glu Thr Gly Arg Lys Leu Thr Pro Lys Glu Ala Phe
228          715          720          725
230 cgg cag ctg tcg cac cgc ttc cat ggc aag ggc tca ggc aag atg aag 2264
231 Arg Gln Leu Ser His Arg Phe His Gly Lys Gly Ser Gly Lys Met Lys
232          730          735          740
234 aca gag cgg cgg atg aag aag ctg gac gag gag gcg ctc ctg aag aag 2312
235 Thr Glu Arg Arg Met Lys Lys Leu Asp Glu Glu Ala Leu Leu Lys Lys
236          745          750          755
238 atg agc tcc agc gac acg ccc ctg ggc acc gtg gcc ctg ctc cag gag 2360
239 Met Ser Ser Ser Asp Thr Pro Leu Gly Thr Val Ala Leu Leu Gln Glu
240          760          765          770
242 aag cag aag gct cag aag acc ccc tac atc gtg ctc agc ggc agc ggc 2408
243 Lys Gln Lys Ala Gln Lys Thr Pro Tyr Ile Val Leu Ser Gly Ser Gly
244          775          780          785          790
246 aag agc atg aac gcg aac acc atc acc aag tgacagcgcc ctcccgtagt 2458
247 Lys Ser Met Asn Ala Asn Thr Ile Thr Lys
248          795          800
250 cggccctgcc tcaaccttca tattaataaa agctccctcc ttatttttaa aaaaaaaaaa 2518
252 aaaaaaaaaa 2527
255 <210> SEQ ID NO: 2
256 <211> LENGTH: 800
257 <212> TYPE: PRT
258 <213> ORGANISM: Homo sapiens

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/202,047A

DATE: 05/07/2002

TIME: 15:57:05

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw

260 &lt;400&gt; SEQUENCE: 2

```

261 Met Gly Ser Ser Lys Lys His Arg Gly Glu Lys Glu Ala Ala Gly Thr
262   1           5           10           15
264 Thr Ala Ala Ala Gly Thr Gly Gly Ala Thr Glu Gln Pro Pro Arg His
265           20           25           30
267 Arg Glu His Lys Lys His Lys His Arg Ser Gly Gly Ser Gly Gly Ser
268           35           40           45
270 Gly Gly Glu Arg Arg Lys Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly
271   50           55           60
273 Ser Gly Arg Arg Gly Ala Glu Ala Glu Ala Arg Ser Ser Thr His Gly
274  65           70           75           80
276 Arg Glu Arg Ser Gln Ala Glu Pro Ser Glu Arg Arg Val Lys Arg Glu
277           85           90           95
279 Lys Arg Asp Asp Gly Tyr Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser
280           100          105          110
282 Gly Asp Ala Ser Ser Leu Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala
283           115          120          125
285 Lys Leu Gly Leu Lys Pro Leu Glu Val Asn Ala Ile Lys Lys Glu Ala
286           130          135          140
288 Gly Thr Lys Glu Glu Pro Val Thr Ala Asp Val Ile Asn Pro Met Ala
289 145          150          155          160
291 Leu Arg Gln Arg Glu Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu
292           165          170          175
294 Lys Arg Leu Leu Asn Gln Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu
295           180          185          190
297 Asp Asp Pro Trp Leu Asp Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg
298           195          200          205
300 Gln Leu Gln Lys Glu Lys Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu
301           210          215          220
303 Glu Glu Met Asp Gln Glu Phe Gly Val Ser Thr Leu Val Glu Glu Glu
304 225          230          235          240
306 Phe Gly Gln Arg Arg Gln Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly
307           245          250          255
309 Leu Thr Val Glu His Ala Ile Asp Ser Phe Arg Glu Gly Glu Thr Met
310           260          265          270
312 Ile Leu Thr Leu Lys Asp Lys Gly Val Leu Gln Glu Glu Glu Asp Val
313           275          280          285
315 Leu Val Asn Val Asn Leu Val Asp Lys Glu Arg Ala Glu Lys Asn Val
316           290          295          300
318 Glu Leu Arg Lys Lys Lys Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu
319 305          310          315          320
321 Ser Val Asp Asp Leu Ala Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys
322           325          330          335
324 Tyr Asp Glu Glu Leu Glu Gly Glu Arg Pro His Ser Phe Arg Leu Glu
325           340          345          350
327 Gln Gly Gly Thr Ala Asp Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile
328           355          360          365
330 Arg Ala Lys Leu Arg Leu Gln Ala Gln Ser Leu Ser Thr Val Gly Pro
331           370          375          380

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/202,047A

DATE: 05/07/2002

TIME: 15:57:06

Input Set : A:\2002-04-22 0020-4491P seq list.txt

Output Set: N:\CRF3\05072002\I202047A.raw